

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: DOMPE' S.p.A.
- (B) STREET: Via Campo di Pile
- (C) CITY: L'AQUILA
- (E) COUNTRY: ITALY
- (F) POSTAL CODE (ZIP): 67100

(ii) TITLE OF INVENTION: RECOMBINANT PROTEINS DERIVED FROM HGF AND MSP

(iii) NUMBER OF SEQUENCES: 4

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1725 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGTGGGTGA	CCAAACTCCT	GCCAGCCCTG	CTGCTGCAGC	ATGTCCCTCT	GCATCTCCCT	60
CTGCTCCCCA	TCGCCATCCC	CTATGCAGAG	GGACAAAAGGA	AAAGAAGAAA	TACAATTCAT	120
GAATTCAAAA	AATCAGCAAA	GACTACCCTA	ATCAAAATAG	ATCCAGCACT	GAAGATAAAA	180
ACCAAAAAAG	TGAATACTGC	AGACCAATGT	GCTAATAGAT	GTACTAGGAA	TAAAGGACTT	240
CCATTCACTT	GCAAGGCTTT	TGTTTTGAT	AAAGCAAGAA	AAACATGCCT	CTGGTTCCCC	300
TTCAATAGCA	TGTCAAGTGG	AGTGAAAAAA	GAATTGGCC	ATGAATTGAA	CCTCTATGAA	360
AACAAAGACT	ACATTAGAAA	CTGCATCATT	GGTAAAGGAC	GCAGCTACAA	GGGAACAGTA	420
TCTATCACTA	AGAGTGGCAT	CAAATGTCAG	CCCTGGAGTT	CCATGATACC	ACACGAACAC	480
AGCTATCGGG	GTAAAGACCT	ACAGGAAAAC	TACTGTCGAA	ATCCTCGAGG	GGAAGAAGGG	540
GGACCCTGGT	GTTCACAAG	CAATCCAGAG	GTACGCTACG	AACTCTGTGA	CATTCCCTCAG	600

*sub
B11-1
Cont*

TGNTCAGAAG TTGAATGCAT GACCTGCAAT GGGGAGAGTT ATCGAGGTCT CATGGATCAT 660
ACAGAACATGAG GCAAGATTTG TCAGCGCTGG GATCATCAGA CACCACACCG GCACAAATTG 720
TTGCCTGAAA GATATCCCGA CAAGGGCTTT GATGATAATT ATTGCCGCAA TCCCGATGGC 780
CAGCCGAGGC CATGGTGCTA TACTCTTGAC CCTCACACCC GCTGGGAGTA CTGTGCAATT 840
AAAACATGCG CTGACAAAGC TTCGGGCGGT GGCGGTTCTG GTGGCGGTGG CTCCGGCGGT 900
GGCGGTTCTC TAGAGGGACA AAGGAAAAGA AGAAATACAA TTCATGAATT CAAAAAATCA 960
GCAAAGACTA CCCTAATCAA AATAGATCCA GCACTGAAGA TAAAAACCAA AAAAGTGAAT 1020
ACTGCAGACC AATGTGCTAA TAGATGTACT AGGAATAAAG GACTTCCATT CACTTGCAAG 1080
GCTTTGTTT TTGATAAAGC AAGAAAACAA TGCCCTCTGGT TCCCCTTCAA TAGCATGTCA 1140
AGTGGAGTGA AAAAAGAATT TGGCCATGAA TTTGACCTCT ATGAAAACAA AGACTACATT 1200
AGAAAATGCA TCATTGGTAA AGGACGCAGC TACAAGGGAA CAGTATCTAT CACTAAGAGT 1260
GGCATCAAAT GTCAGCCCTG GAGTTCCATG ATACCACACG AACACAGCTA TCGGGGTAAA 1320
GACCTACAGG AAAACTACTG TCGAAATCCT CGAGGGGAAG AAGGGGGACC CTGGTGTTTC 1380
ACAAGCAATC CAGAGGTACG CTACGAAGTC TGTGACATTC CTCAGTGTTC AGAAGTTGAA 1440
TGCATGACCT GCAATGGGGA GAGTTATCGA GGTCTCATGG ATCATAACAGA ATCAGGCAAG 1500
ATTGTCAGC GCTGGGATCA TCAGACACCA CACCGGCACA AATTCTTGCC TGAAAGATAT 1560
CCCGACAAGG GCTTTGATGA TAATTATTGC CGCAATCCCG ATGGCCAGCC GAGGCCATGG 1620
TGCTATACTC TTGACCCCTCA CACCCGCTGG GAGTACTGTG CAATTAAAAC ATGCGCTGAC 1680
AAAGCTGACG ACGACGACAA ACACCAACAC CACCAACACC ACTAG 1725

DOCUMENT-09425026

3

(2) INFORMATION FOR SEQ ID NO: 2:

*Sub
B11
Cont'd*

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 574 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met	Trp	Val	Thr	Lys	Leu	Leu	Pro	Ala	Leu	Leu	Leu	Gln	His	Val	Leu
1				5					10				15		

Leu	His	Leu	Leu	Leu	Leu	Pro	Ile	Ala	Ile	Pro	Tyr	Ala	Glu	Gly	Gln
							20		25			30			

Arg	Lys	Arg	Arg	Asn	Thr	Ile	His	Glu	Phe	Lys	Lys	Ser	Ala	Lys	Thr
					35				40			45			

Thr	Leu	Ile	Lys	Ile	Asp	Pro	Ala	Leu	Lys	Ile	Lys	Thr	Lys	Lys	Val
					50				55			60			

Asn	Thr	Ala	Asp	Gln	Cys	Ala	Asn	Arg	Cys	Thr	Arg	Asn	Lys	Gly	Leu
					65			70		75		80			

Pro	Phe	Thr	Cys	Lys	Ala	Phe	Val	Phe	Asp	Lys	Ala	Arg	Lys	Gln	Cys
					85			90			95				

Leu	Trp	Phe	Pro	Phe	Asn	Ser	Met	Ser	Ser	Gly	Val	Lys	Lys	Glu	Phe
					100			105			110				

Gly	His	Glu	Phe	Asp	Leu	Tyr		Glu	Asn	Lys	Asp	Tyr	Ile	Arg	Asn	Cys
					115			120			125					

Ile	Ile	Gly	Lys	Gly	Arg	Ser	Tyr	Lys	Gly	Thr	Val	Ser	Ile	Thr	Lys
					130			135			140				

Ser	Gly	Ile	Lys	Cys	Gln	Pro	Trp	Ser	Ser	Met	Ile	Pro	His	Glu	His
					145			150			155			160	

Ser	Tyr	Arg	Gly	Lys	Asp	Leu	Gln	Glu	Asn	Tyr	Cys	Arg	Asn	Pro	Arg
					165			170			175				

Gly	Glu	Gly	Gly	Pro	Trp	Cys	Phe	Thr	Ser	Asn	Pro	Glu	Val	Arg
					180			185			190			

Tyr	Glu	Val	Cys	Asp	Ile	Pro	Gln	Cys	Ser	Glu	Val	Glu	Cys	Met	Thr
					195			200			205				

Cys	Asn	Gly	Glu	Ser	Tyr	Arg	Gly	Leu	Met	Asp	His	Thr	Glu	Ser	Gly
					210			215			220				

Sub
 B11
 Cont'd

Lys	Ile	Cys	Gln	Arg	Trp	Asp	His	Gln	Thr	Pro	His	Arg	His	Lys	Phe
225					230					235					240
Leu	Pro	Glu	Arg	Tyr	Pro	Asp	Lys	Gly	Phe	Asp	Asp	Asn	Tyr	Cys	Arg
				245				250					255		
Asn	Pro	Asp	Gly	Gln	Pro	Arg	Pro	Trp	Cys	Tyr	Thr	Leu	Asp	Pro	His
				260			265				270				
Thr	Arg	Trp	Glu	Tyr	Cys	Ala	Ile	Lys	Thr	Cys	Ala	Asp	Lys	Ala	Ser
			275				280				285				
Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Ser	Leu		
			290		295				300						
Glu	Gly	Gln	Arg	Lys	Arg	Arg	Asn	Thr	Ile	His	Glu	Phe	Lys	Lys	Ser
		305			310					315					320
Ala	Lys	Thr	Thr	Leu	Ile	Lys	Ile	Asp	Pro	Ala	Leu	Lys	Ile	Lys	Thr
				325				330			335				
Lys	Lys	Val	Asn	Thr	Ala	Asp	Gln	Cys	Ala	Asn	Arg	Cys	Thr	Arg	Asn
				340				345			350				
Lys	Gly	Leu	Pro	Phe	Thr	Cys	Lys	Ala	Phe	Val	Phe	Asp	Lys	Ala	Arg
				355			360			365					
Lys	Gln	Cys	Leu	Trp	Phe	Pro	Phe	Asn	Ser	Met	Ser	Ser	Gly	Val	Lys
				370			375			380					
Lys	Glu	Phe	Gly	His	Glu	Phe	Asp	Leu	Tyr	Glu	Asn	Lys	Asp	Tyr	Ile
				385		390				395					400
Arg	Asn	Cys	Ile	Ile	Gly	Lys	Gly	Arg	Ser	Tyr	Lys	Gly	Thr	Val	Ser
				405				410			415				
Ile	Thr	Lys	Ser	Gly	Ile	Lys	Cys	Gln	Pro	Trp	Ser	Ser	Met	Ile	Pro
				420				425			430				
His	Glu	His	Ser	Tyr	Arg	Gly	Lys	Asp	Leu	Gln	Glu	Asn	Tyr	Cys	Arg
				435			440				445				
Asn	Pro	Arg	Gly	Glu	Glu	Gly	Gly	Pro	Trp	Cys	Phe	Thr	Ser	Asn	Pro
				450		455				460					
Glu	Val	Arg	Tyr	Glu	Val	Cys	Asp	Ile	Pro	Gln	Cys	Ser	Glu	Val	Glu
				465		470				475					480
Cys	Met	Thr	Cys	Asn	Gly	Glu	Ser	Tyr	Arg	Gly	Leu	Met	Asp	His	Thr
				485				490			495				
Glu	Ser	Gly	Lys	Ile	Cys	Gln	Arg	Trp	Asp	His	Gln	Thr	Pro	His	Arg
				500				505			510				
His	Lys	Phe	Leu	Pro	Glu	Arg	Tyr	Pro	Asp	Lys	Gly	Phe	Asp	Asp	Asn
				515			520			525					

5

Tyr	Cys	Arg	Asn	Pro	Asp	Gly	Gln	Pro	Arg	Pro	Trp	Cys	Tyr	Thr	Leu
530							535					540			
Asp	Pro	His	Thr	Arg	Trp	Glu	Tyr	Cys	Ala	Ile	Lys	Thr	Cys	Ala	Asp
545						550				555			560		
Lys	Ala	Asp	Asp	Asp	Asp	Lys	His								
						565					570				

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1692 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATGGGGTGGC	TCCCACTCCT	GCTGCTTCTG	ACTCAATGCT	TAGGGGTCCC	TGGGCAGCGC	60
TCGCCATTGA	ATGACTTCCA	AGTGCTCCGG	GGCACAGAGC	TACAGCACCT	GCTACATGCG	120
GTGGTGCCCC	GGCCTTGGCA	GGAGGATGTG	GCAGATGCTG	AAGAGTGTGC	TGGTCGCTGT	180
GGGCCCTTAA	TGGACTGCCG	GGCCTTCCAC	TACAACGTGA	GCAGCCATGG	TTGCCAACTG	240
CTGCCATGGA	CTCAACACTC	GCCCCACACG	AGGCTGCGGC	GTTCTGGCG	CTGTGACCTC	300
TTCCAGAAGA	AAGACTACGT	ACGGACCTGC	ATCATGAACA	ATGGGGTTGG	GTACCGGGC	360
ACCATGGCCA	CGACCGTGGG	TGGCCTGCC	TGCCAGGCTT	GGAGCCACAA	GTTCCCGAAT	420
GATCACAAAGT	ACACGCCAC	TCTCCGGAAT	GGCCTGGAAG	AGAACTTCTG	CCGTAACCCT	480
GATGGCGACC	CCGGAGGTCC	TTGGTGTAC	ACAAACAGACC	CTGCTGTGCG	CTTCCAGAGC	540
TGCGGCATCA	AATCCTGCCG	GGAGGCCGCG	TGTGTCTGGT	GCAATGGCGA	GGAATACCGC	600
GGCGCGGTAG	ACCGCACCGA	GTCAGGGCGC	GAGTGCCAGC	GCTGGGATCT	TCAGCACCCG	660
CACCAGCACC	CCTTCGAGCC	GGGCAAGTTC	CTCGACCAAG	GTCTGGACGA	CAACTATTGC	720
CGGAATCCTG	ACGGCTCCGA	GCGGCCATGG	TGCTACACTA	CGGATCCGCA	GATCGAGCGA	780
GAGTTCTGTG	ACCTCCCCCG	CTGCGGGTCC	GAGGCACAGC	CCCGCCTCGA	GGGCGGTGGC	840
GGTTCTGGTG	GCGGTGGCTC	CGGCAGGTGGC	GGTTCTCTAG	AGGGACAAAG	GAAAAGAAGA	900
AAATACAATTC	ATGAATTCAA	AAAATCAGCA	AAGACTACCC	TAACTAAAAT	AGATCCAGCA	960

CTGAAGATAA AAACCAAAAA AGTGAATACT GCAGACCAAT GTGCTAATAG ATGTACTAGG 1020
 AATAAAGGAC TTCCATTACAC TTGCAAGGCT TTTGTTTTG ATAAAGCAAG AAAACAATGC 1080
 CTCTGGTTCC CCTTCATAG CATGTCAAGT GGAGTGAAAA AAGAATTGG CCATGAATTT 1140
 GACCTCTATG AAAACAAAGA CTACATTAGA AACTGCATCA TTGGTAAAGG ACGCAGCTAC 1200
 AAGGGAACAG TATCTATCAC TAAGAGTGGC ATCAAATGTC AGCCCTGGAG TTCCATGATA 1260
 CCACACGAAC ACAGCTATCG GGGTAAAGAC CTACAGGAAA ACTACTGTCG AAATCCTCGA 1320
 GGGGAAGAAG GGGGACCCTG GTGTTTCACA AGCAATCCAG AGGTACGCTA CGAAGTCTGT 1380
 GACATTCCCTC AGTGTTCAGA AGTTGAATGC ATGACCTGCA ATGGGGAGAG TTATCGAGGT 1440
 CTCATGGATC ATACAGAATC AGGCAAGATT TGTCAGCGCT GGGATCATCA GACACCACAC 1500
 CGGCACAAAT TCTTGCCTGA AAGATATCCC GACAAGGGCT TTGATGATAA TTATTGCCGC 1560
 AATCCCGATG GCCAGCCGAG GCCATGGTGC TATACTCTTG ACCCTCACAC CCGCTGGAG 1620
 TACTGTGCAA TTAAAACATG CGCTGACAAA GCTGACGACG ACGACAAACA CCACCACAC 1680
 CACCACCACT AG 1692

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 563 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met	Gly	Trp	Leu	Pro	Leu	Leu	Leu	Leu	Thr	Gln	Cys	Leu	Gly	Val
1														15

Pro	Gly	Gln	Arg	Ser	Pro	Leu	Asn	Asp	Phe	Gln	Val	Leu	Arg	Gly	Thr
															30
20															

Glu	Leu	Gln	His	Leu	Leu	His	Ala	Val	Val	Pro	Gly	Pro	Trp	Gln	Glu
															45
35															

Asp	Val	Ala	Asp	Ala	Glu	Glu	Cys	Ala	Gly	Arg	Cys	Gly	Pro	Leu	Met
															60
50															

Asp	Cys	Arg	Ala	Phe	His	Tyr	Asn	Val	Ser	Ser	His	Gly	Cys	Gln	Leu
															80
65															
															75

DOCUMENT EDITION 001

pub
bill
Cont'd

Leu Pro Trp Thr Gln His Ser Pro His Thr Arg Leu Arg Arg Ser Gly
85 90 95

*sub
B1
cont'd*

Arg Cys Asp Leu Phe Gln Lys Lys Asp Tyr Val Arg Thr Cys Ile Met
100 105 110

Asn Asn Gly Val Gly Tyr Arg Gly Thr Met Ala Thr Thr Val Gly Gly
115 120 125

Leu Pro Cys Gln Ala Trp Ser His Lys Phe Pro Asn Asp His Lys Tyr
130 135 140

Thr Pro Thr Leu Arg Asn Gly Leu Glu Glu Asn Phe Cys Arg Asn Pro
145 150 155 160

Asp Gly Asp Pro Gly Gly Pro Trp Cys Tyr Thr Thr Asp Pro Ala Val
165 170 175

Arg Phe Gln Ser Cys Gly Ile Lys Ser Cys Arg Glu Ala Ala Cys Val
180 185 190

Trp Cys Asn Gly Glu Glu Tyr Arg Gly Ala Val Asp Arg Thr Glu Ser
195 200 205

Gly Arg Glu Cys Gln Arg Trp Asp Leu Gln His Pro His Gln His Pro
210 215 220

Phe Glu Pro Gly Lys Phe Leu Asp Gln Gly Leu Asp Asp Asn Tyr Cys
225 230 235 240

Arg Asn Pro Asp Gly Ser Glu Arg Pro Trp Cys Tyr Thr Thr Asp Pro
245 250 255

Gln Ile Glu Arg Glu Phe Cys Asp Leu Pro Arg Cys Gly Ser Glu Ala
260 265 270

Gln Pro Arg Leu Glu Gly Gly Ser Gly Gly Ser Gly
275 280 285

Gly Gly Ser Leu Glu Gly Gln Arg Lys Arg Arg Asn Thr Ile His
290 295 300

Glu Phe Lys Lys Ser Ala Lys Thr Thr Leu Ile Lys Ile Asp Pro Ala
305 310 315 320

Leu Lys Ile Lys Thr Lys Lys Val Asn Thr Ala Asp Gln Cys Ala Asn
325 330 335

Arg Cys Thr Arg Asn Lys Gly Leu Pro Phe Thr Cys Lys Ala Phe Val
340 345 350

Phe Asp Lys Ala Arg Lys Gln Cys Leu Trp Phe Pro Phe Asn Ser Met
355 360 365

Ser Ser Gly Val Lys Lys Glu Phe Gly His Glu Phe Asp Leu Tyr Glu
370 375 380

Asn Lys Asp Tyr Ile Arg Asn Cys Ile Ile Gly Lys Gly Arg Ser Tyr
385 390 395 400

Lys Gly Thr Val Ser Ile Thr Lys Ser Gly Ile Lys Cys Gln Pro Trp
405 410 415

Ser Ser Met Ile Pro His Glu His Ser Tyr Arg Gly Lys Asp Leu Gln
420 425 430

Glu Asn Tyr Cys Arg Asn Pro Arg Gly Glu Glu Gly Gly Pro Trp Cys
435 440 445

Phe Thr Ser Asn Pro Glu Val Arg Tyr Glu Val Cys Asp Ile Pro Gln
450 455 460

Cys Ser Glu Val Glu Cys Met Thr Cys Asn Gly Glu Ser Tyr Arg Gly
465 470 475 480

Leu Met Asp His Thr Glu Ser Gly Lys Ile Cys Gln Arg Trp Asp His
485 490 495

Gln Thr Pro His Arg His Lys Phe Leu Pro Glu Arg Tyr Pro Asp Lys
500 505 510

Gly Phe Asp Asp Asn Tyr Cys Arg Asn Pro Asp Gly Gln Pro Arg Pro
515 520 525

Trp Cys Tyr Thr Leu Asp Pro His Thr Arg Trp Glu Tyr Cys Ala Ile
530 535 540

Lys Thr Cys Ala Asp Lys Ala Asp Asp Asp Asp Lys His His His His
545 550 555 560

His His His

DOCUMENT RELEASED UNDER E.O. 14176

MW
B II
cont